A Caged Lanthanide Complex as a Paramagnetic Shift Agent for Protein NMR

Miguel Prudêncio,[a] Jan Rohovec,[b] Joop A. Peters,[b] Elitza Tocheva,[c] Martin J. Boulanger,[c] Michael E. P. Murphy,[c] Hermen-Jan Hupkes,[d] Walter Kosters,[d] Antonietta Impagliazzo,[a] and Marcellus Ubbink*[a]

Abstract: A lanthanide complex, named CLaNP (caged lanthanide NMR probe) has been developed for the characterisation of proteins by paramagnetic NMR spectroscopy. The probe consists of a lanthanide chelated by a derivative of DTPA (diethylenetriaminepentaacetic acid) with two thiol reactive functional groups. The CLaNP molecule is attached to a protein by two engineered, surface-exposed, Cys residues in a bidentate manner. This drastically limits the dynamics of the metal relative to the protein and enables measurements of pseudocontact shifts. NMR spectroscopy experiments on a diamagnetic control and the crystal structure of the probe-protein complex demonstrate that the protein structure is not affected by probe attachment. The probe is able to induce pseudocontact shifts to at least 40 Å from the metal and causes residual dipolar couplings due to alignment at a high magnetic field. The molecule exists in several isomeric forms with different paramagnetic tensors; this provides a fast way to obtain long-range distance restraints.

Keywords: cage compounds · lanthanides · NMR spectroscopy · proteins · pseudocontact shift

Introduction

Unpaired electrons present in radicals and metal ions cause paramagnetic effects that strongly influence NMR spectra. For some time, paramagnetism was considered a nuisance in NMR structure determination.[1,2] However, in recent years, it has been demonstrated that paramagnetism can provide useful structural information and, as a result, determination of the solution structure of proteins containing paramagnetic metals is now a reality.[3,4] The presence of an unpaired electron can cause both broadening and shifts of the resonances in the NMR spectrum. Line broadening is a consequence of enhanced relaxation. To a good approximation, this effect is isotropic and falls off with the sixth power of the distance between the metal and the observed nucleus. It is, therefore, a sensitive tool for distance measurements, but it is localised to the immediate surroundings of the metal or radical.[1] On the other hand, line shifts are caused by two effects: a) contact shifts are due to delocalisation of the unpaired electron onto the nucleus and are only relevant for nuclei not more than several bonds away from the metal and b) pseudocontact shifts (PCS) result from the dipolar interaction between the (time averaged) unpaired spin and the nucleus, and their size depends on the anisotropy of the magnetic susceptibility.[1] The magnitude of PCS is described by the magnetic susceptibility tensor ([Eq. (1)] in the Experimental Section) and falls off with the third power of the distance between the metal and the observed nucleus. For strongly paramagnetic metals with a high anisotropy, PCS can be observed for nuclei at long distances from the metal,[5,6] which can be used to provide restraints for structure determination.[2,4] Paramagnetic molecules with strongly anisotropic magnetic

[a] Dr. M. Prudêncio, A. Impagliazzo, Dr. M. Ubbink
Leiden Institute of Chemistry
Gorlaeus Laboratories, Leiden University
P.O. Box 9502, 2300 RA Leiden (The Netherlands)
Fax: (+31) 71-5274-349
E-mail: m.ubbink@chem.leidenuniv.nl

[b] Dr. J. Rohovec, Dr. J. A. Peters
Laboratory of Applied Organic Chemistry and Catalysis
Delft University of Technology, Julianalaan 136
2628 BL Delft (The Netherlands)

[c] E. Tocheva, Dr. M. J. Boulanger, Dr. M. E. P. Murphy
Department of Microbiology and Immunology
University of British Columbia, #300–6174
University Boulevard, Vancouver BC
V6T 1Z3 (Canada)

[d] H.-J. Hupkes, Dr. W. Kosters
Leiden Institute of Advanced Computer Science
Leiden University, Niels Bohrweg 1
2333 CA Leiden (The Netherlands)

Supporting information for this article is available on the WWW under http://www.chemeurj.org/ or from the author. This includes: Table S1: Assignments of the ^{1}H and ^{15}N amide nuclei in Zn-pseudoazurin. Supplementary experimental methods SM1-Tensor determination.

DOI: 10.1002/chem.200306019

susceptibility also show some degree of spontaneous alignment at high magnetic fields; this provides an easy way to obtain residual dipolar couplings.[7]

To take advantage of these paramagnetic effects in diamagnetic molecules, various types of paramagnetic probe molecules have been used. The first group consists of soluble probes, which are not covalently attached to the molecule of interest, such as paramagnetic metals[8–14] and nitroxide spin labels.[15–17] The structural information provided by this category of probes is limited by the fact that any observed PCS are very small, due to averaging effects and by the limited range of paramagnetic relaxation effects they induce. The second group includes probes that are attached to the molecules of interest by selective covalent or coordination bonds, such as nitroxide spin labels,[18–21] divalent cobalt[22–24] or manganese[25] ions and lanthanides (Ln).[26] Spin labels have a limited action radius and cause no PCS, while Co^{2+} requires a specific binding site and induces a limited range of PCS .

Paramagnetic lanthanides, with their unpaired electrons located in the 4f inner shell, are superior paramagnetic probes, because a) a range of paramagnetic effects can be expected with different lanthanides, b) similar ionic radii and geometries are observed for complexes of different lanthanides and c) lanthanum(III), lutetium(III) and yttrium(III) can be used as diamagnetic analogues. These advantages were demonstrated for the Ca^{2+}-binding protein calbindin, in which the calcium atom was replaced by various lanthanides.[27–29] Obviously, this application of lanthanides is limited to proteins that contain a metal site capable of accommodating a Ln^{3+} ion and assumes that metal replacement is isomorphic. In addition, it results in the broadening of resonances in the vicinity of the Ln^{3+} ion due to relaxation effects.

Recently, various covalent paramagnetic tags have been developed and inserted on a genetic level at one of the termini of the protein.[30–33] Their main purpose is to cause spontaneous alignment to obtain residual dipolar couplings. They are less useful for restraints based on PCS or relaxation, because of limited choice in location and the dynamics of the probe relative to the protein.

We have designed a molecule that can be used as a paramagnetic probe without the limitations previously mentioned. The caged lanthanide NMR probe (CLaNP) can be specifically attached to the surface of a protein where two cysteine residues have been engineered at a convenient distance from each other. This probe constitutes, in effect, an artificial paramagnetic centre that can be positioned specifically on an otherwise diamagnetic protein, without causing significant structural changes. The CLaNP molecule binds to the protein in a bidentate way, thus it has reduced mobility and allows PCS to be measured. Although PCS have been reported once by using a similar probe attached to a single protein residue,[34] it appears that generally the probe is much too dynamic in such cases; this is in accord with the fact that there are six rotatable bonds between the probe and the backbone of the protein. Linking the probe at two positions on the protein surface results in a cyclic compound with severely reduced mobility.

The fact that there are at least eight bonds between the lanthanide ion and the nuclei in the protein means that contact shift effects are negligible. Relaxation-induced line broadening is also negligible or limited to very few residues, because the metal in the probe is located >6 Å away from the nearest amide. Finally, the molecule is able to cause spontaneous partial alignment of the protein at high magnetic fields; this enables residual dipolar couplings to be measured. This newly designed probe is able to accommodate any lanthanide ion and can, in principle, be useful in providing variable-range structural information about large proteins and protein complexes.

Results and Discussion

Design of the ClaNP molecule: The design of a molecule that can be used as an artificial paramagnetic probe, capable of generating useful structural information about proteins, requires a number of conditions to be met. The most important of these are a) that the probe reacts with specific residues in the protein, thus enabling its attachment point to be conveniently selected; b) that it has little flexibility with regard to the protein to prevent averaging of the pseudocontact effect; c) that it can accommodate a paramagnet capable of generating a useful range of paramagnetic effects and d) that its attachment does not significantly alter the structure and stability of the protein.

Cys residues rarely occur on protein surfaces and thiol chemistry can be used for specific attachment reactions. DTPA is an excellent chelator of lanthanides, yielding a soluble complex, which can easily be modified. For these reasons, two thiol reactive functional groups were added to DTPA (Figure 1 top) for the attachment to two engineered Cys residues on a protein surface. Attachment of this modified DTPA–Ln complex in a bidentate manner is aimed at drastically limiting its mobility relative to the protein. DTPA-bis(amides) are known to bind Ln^{3+} ions in a cage-like, octadentate fashion through the three nitrogen atoms of the diethylenetriamine backbone, three carboxylate oxygen atoms, and two amide oxygen atoms; overall this yields neutral complexes.[35] In the case of the DTPA-bis(amide) of S-(2-methylaminomethyl)methanesulfothioate (MTS) described in this work, the introduced sulfothioate groups are expected to be ≈ 10 Å from each other (Figure 1 middle). During the reaction with two Cys residues (Figure 1 bottom), a sulphinic acid molecule leaves each of the two functional groups on the probe molecule and the protein–CLaNP complex is formed.

Lanthanides are the metals of choice, because of their superior paramagnetic properties discussed above. However, a disadvantage is their high coordination number, hence the prevalence of isomeric forms. These complexes have four chiral centres; the three nitrogen atoms of the diethylenetriamine moiety and the Ln^{3+} ion. Consequently, they may occur in up to 16 ($=2^4$) enantiomeric forms. Upon binding of the probe molecule to the protein, the DTPA-bis(amide) becomes part of a cyclic structure and a maximum number of eight diastereomers with different exchange rates may be

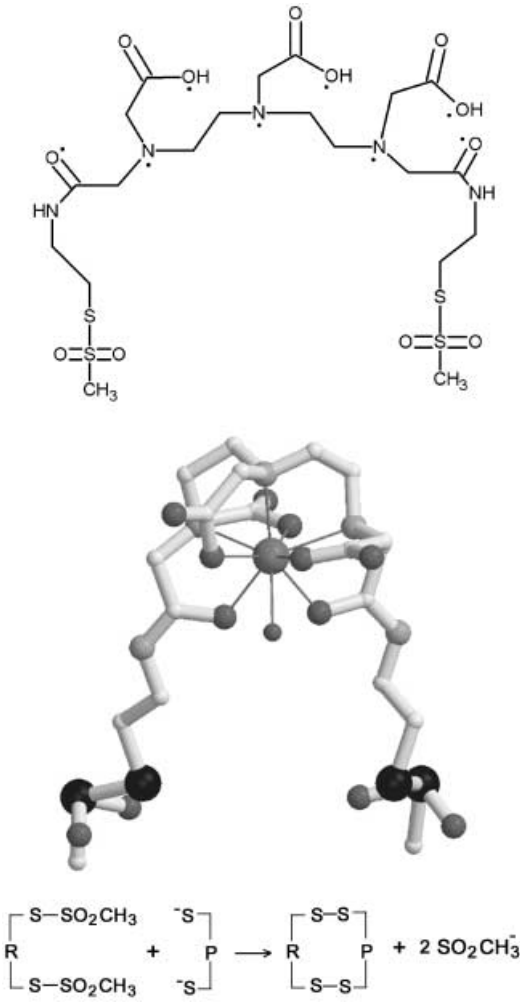

Figure 1. The probe molecule. Top; chemical structure of the bis(MTS) derivative of DTPA. The dots represent the atoms involved in the coordination of the lanthanide ion. Middle; ball-and-stick representation of the Ln-(bis(MTS)–DTPA) complex, with carbon, nitrogen and oxygen atoms in white, light grey and dark grey, respectively. Sulphur atoms are represented by large black spheres and the water oxygen atom and Ln atoms are shown as small and a large dark grey spheres, respectively. Bottom; schematic representation of the reaction between the probe molecule $[R(S_2O_2CH_3)_2]$ and two thiol groups on a protein, P.

envisaged; this has consequences for the paramagnetic effects (see below).

Construction of the CLaNP–protein complex: To characterise its paramagnetic properties, the CLaNP molecule was attached to pseudoazurin (Psaz) from *Alcaligenes faecalis* S6, a well studied copper-containing electron transfer protein, which does not have any surface-exposed Cys residues. Its size (14 kDa) allows easy NMR characterisation, whilst allowing paramagnetic effects to be measured up to $\approx$40 Å from the metal in the probe. To avoid any interference with sulphur chemistry and to exclude undesired paramagnetic effects, the copper in this protein was replaced by zinc, a change that is known to have a negligible effect on the structure of type 1 Cu proteins.[36]

The distance between the sulphur atoms in the two engineered Cys residues was aimed to be 8–10 Å, to enable attachment of the bidentate probe and to avoid intramolecular sulphur bridge formation. From the numerous residue pairs that fit these criteria, the E51C/E54C double mutant (dCPsaz) was selected and produced. Thiol determination indicated an accessible thiol content of 1.7 ($\pm$0.2) for dCPsaz after reduction with DTT. To confirm the accessibility of the thiol groups, the protein was reacted with MTS (the functional molecule without Ln–DTPA); this resulted in an increase of the mass by 151 Da ($\pm$2) (from 13473 Da to 13624 Da). This confirms the accessibility of the thiol groups in both engineered Cys residues of dCPsaz, because the expected mass increase after reaction with two MTS molecules is 150 Da.

The reduced form of the protein was reacted with either the Y-containing, diamagnetic probe or the Yb-containing, paramagnetic probe and separated from a fraction of dimeric protein present after reaction by gel filtration. Mass spectrometry results revealed a mass of 14068 Da ($\pm$2) for the diamagnetic complex (hereafter ^{Y}CLaNP–dCPsaz) and a mass of 14152 Da ($\pm$2) for the paramagnetic one (hereafter ^{Yb}CLaNP–dCPsaz). Taking into account the mass of both the unreacted diamagnetic and paramagnetic probes (755 Da and 839 Da, respectively), these results indicate that both leaving groups are absent in the complexes, which shows that each of the probes binds dCPsaz by its two reactive arms. Complexes with masses corresponding to the ClaNP molecule being attached through a single disulphide bridge were not observed.

The effect of the mutations and probe binding on the three-dimensional structure of the protein was investigated by comparing the [^{15}N,^{1}H]-HSQC spectra of the diamagnetic ^{Y}CLaNP–dCPsaz complex and of the wild-type ZnPsaz. A plot of the ^{1}H-amide chemical shift difference (^{1}H-δ^{diff}) between the two spectra (Figure 2) shows that the overall structure of ZnPsaz does not change significantly due to the mutations and probe binding.

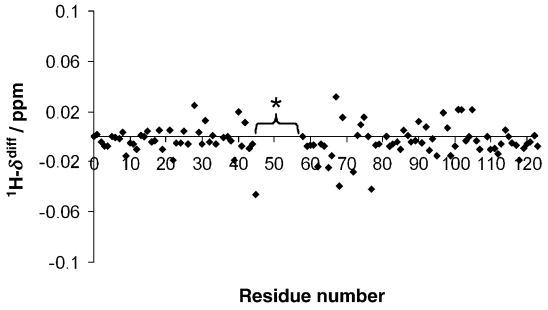

Figure 2. Chemical shift differences between amide protons (^{1}H-δ^{diff}) of wild-type Zn-Psaz and ^{Y}CLaNP–dCPsaz. Residues 46 to 57 (marked with *) are not observed in the spectrum.

The average ^{1}H-amide chemical shift difference observed is less than 0.01 ppm and the overall similarity of the two structures is confirmed by X-ray crystallography results (see below). It should be noted that the resonances corresponding to Cys51 and Cys54 and residues in their immediate vicinity are not observed in the spectrum of the ^{Y}CLaNP–dCPsaz complex. The reason for the disappearance of these

signals is unclear, but it could be caused by line broadening due to an exchange process, most likely within the probe molecule.

X-ray structure: The crystal structure of dCPsaz complexed with the Y-ClaNP probe (PDB entry 1PY0) clearly defines the location of the Y-probe and the Psaz fold (Figure 3 top).

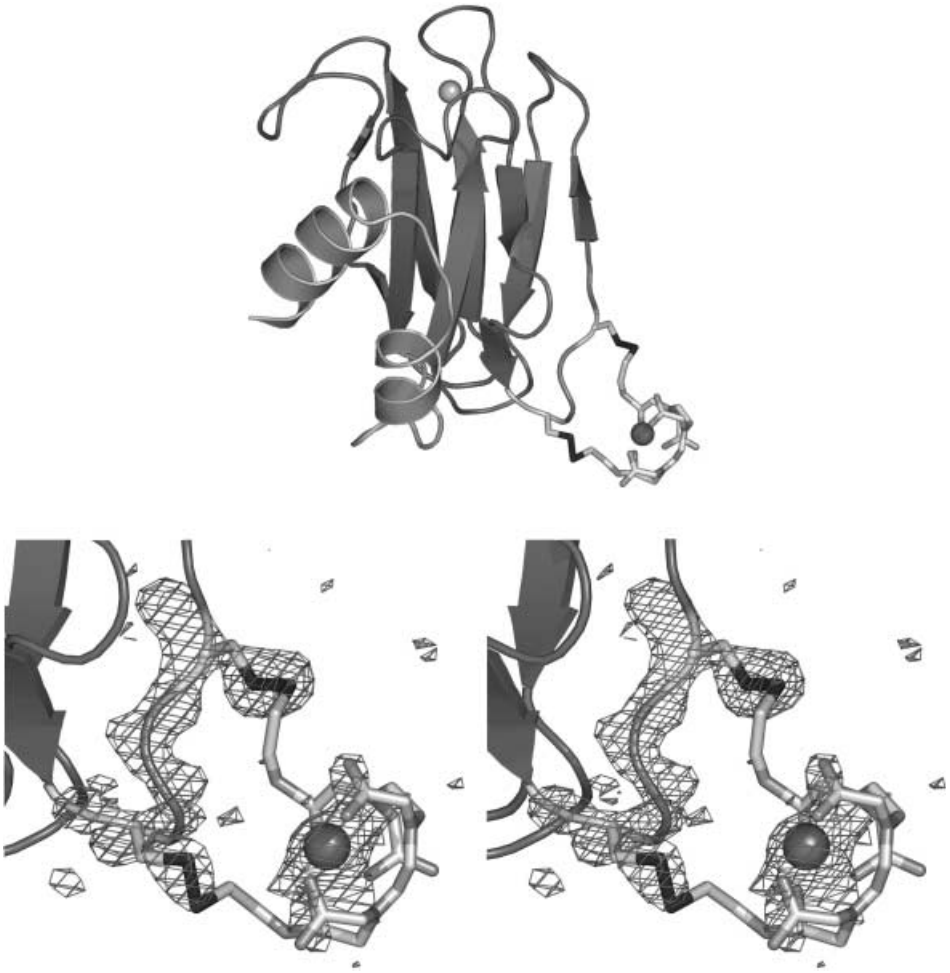

Figure 3. ^{Y}ClaNP–dCPsaz complex crystal structure. Top; the secondary structure elements are coloured as follows: β strands in dark grey and α helices in light grey. The side chains of residues Cys51 and Cys54 and the CLaNP are detailed as sticks. The disulphide bridges between Cys51 and Cys54 to the probe are shown in black. The zinc and the yttrium atoms are depicted as light grey and dark grey spheres, respectively. Bottom; a stereoview of the ^{Y}CLaNP region. A SigmaA[51] weighted difference electron density map of the probe is shown, with residues Cys51 to Cys54 and the yttrium atom contoured at 2.5σ. In both panels, atoms are shown as follows: carbons are white, oxygens and nitrogens are grey and sulphur atoms are in black.

Following structure solution by molecular replacement with an apo-pseudoazurin model, a difference map was computed to locate metals absent in the coordinate set. The highest positive peak in a difference map appeared at the type I site and was modelled as a zinc atom with full occupancy. The next two highest peaks were located in the solvent channels of the crystal. One of these peaks is near the molecular surface and was modelled as a sulphate anion and the other, 6–8 Å from residues Cys51 and Cys54 was modelled as a Y atom. In difference maps of the probe at the end of refinement, density was present for the Y atom, the sulphur linkers and parts of the cage structure (Figure 3 bottom).

The yttrium atom and the ClaNP were refined at 70% occupancy yielding an average B-factor of 68 Å^2. The arms of the probe are poorly defined in the electron density map and the cage structure likely adopts multiple conformations in the crystal. The probe arms are bent such that the cage structure approaches the Psaz peptide. The closest atom of the main chain of Psaz to the yttrium atom is the C$^{\alpha}$ of Gly52, situated 6 Å away. The Y atom is located at 6.0 Å and 8.3 Å from the sulphur atoms of Cys51 and Cys54, respectively and ~30 Å from the zinc atom at the type I site. Overall, there is no change in the fold of the E51C/E54C-ZnPsaz variant structure relative to the native Psaz structure as revealed by an rms deviation between the C$^{\alpha}$ chains of less than 0.3 Å.

Paramagnetic effects: Binding of the paramagnetic probe to dCPsaz causes paramagnetism-induced shifts in the resonances of all amides, as shown by the comparison of the HSQC (heteronuclear single-quantum coherence) spectra of the paramagnetic ^{Yb}CLaNP–dCPsaz complex with that of the diamagnetic control complex, ^{Y}CLaNP–dCPsaz (Figure 4).

Several shifted resonances are observed for each residue as a result of the ability of the CLaNP molecule to isomerise. Our results show that the various observed isomers have different paramagnetic anisotropies and tensor orientations, which result in several resonances for each amide group. The resonances for an amide group approximately lie on a line with a slope of 1. The reason for this is that all PCS, defined in ppm, are independent of the nature of the nucleus that experiences the paramagnetic effect and depend only on its position relative to the metal [Eq. (1) see later]. Amide protons and nitrogens of an amino acid residue are in close proximity of each other and should, therefore, experience similar PCS. Thus, the fact that the shifted peaks lie nearly on a straight line with a slope of 1 represents evidence that the shifts are caused by the pseudocontact effect (Figure 4 bottom). The magnitude of the observed PCS is clearly dependent on the distance between the affected amide and the

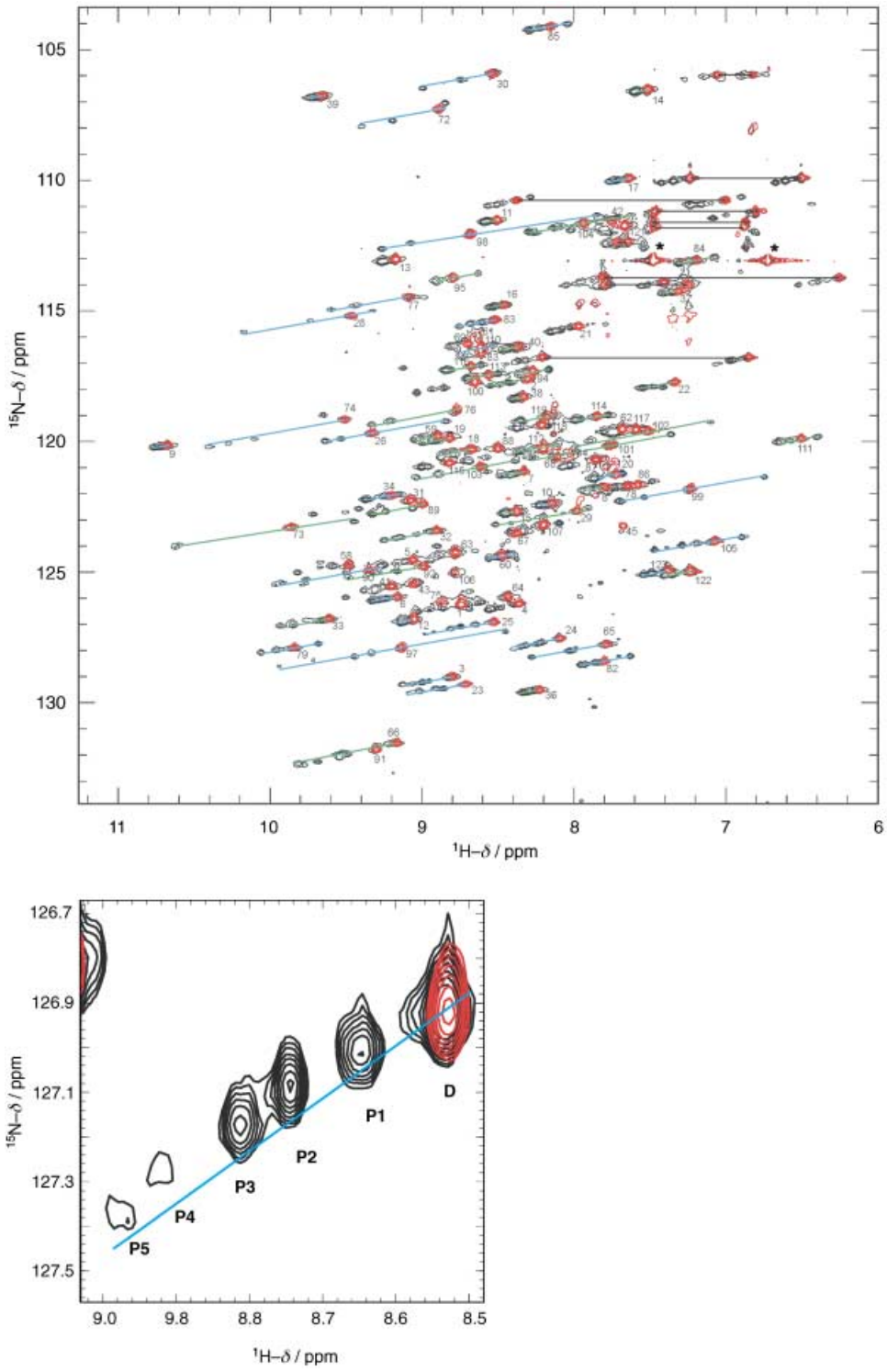

Figure 4. Overlay of the [^{15}N,^{1}H]-HSQC spectra of the ^{Y}CLaNP–dCPsaz (red) and the ^{Yb}CLaNP–dCPsaz (black) complexes. Top; the complete spectrum. The resonances of the 31 amides used as inputs for the determination of the paramagnetic tensors and those of the 37 additional amides assigned with the aid of the determined tensors (see text) are connected by blue and green lines, respectively, with slope 1. Side-chain amide resonances are connected by horizontal black lines. The numeric labels represent the assignments of the protein residues in the diamagnetic complex. The asterisks denote the peaks belonging to ^{15}N-acetamide ($CH_3CO^{15}NH_2$), used as an internal reference. Bottom; the set of resonances for the amide of Ala25. D-diamagnetic resonance; P_1, P_2, P_3, P_4 and P_5- paramagnetically shifted resonances.

Yb atom, in accordance with the third power distance dependence [Eq. (1) see later]. Up to five shifted resonances are observed for each residue, three equally intense ones and two of lower intensity; this probably reflects preferred conformations of the CLaNP molecule. A diamagnetic resonance is also observed for every amide, which corresponds to the fraction of unlabelled dCPsaz in the sample.

Measurements at 17.6 Tesla (750 MHz) of the coupling between amide ^{1}H and ^{15}N atoms showed differences between the paramagnetic and the diamagnetic complexes of up to 10 Hz, which indicate that the ^{Yb}CLaNP–dCPsaz is partially aligned, causing residual dipolar couplings to be observed. Further investigations on partial alignment with other lanthanides are underway.

As expected, for none of the amides that could be observed in the diamagnetic complex, significant line broadening was observed in the paramagnetic complex; this is in accord with the relaxation effect being limited to the immediate surroundings of the lanthanide.

Tensor calculations: From a set of PCS, the size and orientation of the corresponding magnetic susceptibility tensor can be determined, which can then be used for further structural characterisation of a protein or protein complex.[37] However, the presence of multiple paramagnetically shifted resonances for each of the observed amides somewhat complicates tensor determinations, since there is no unequivocal way of mapping the resonances of similar intensity to each tensor. In order to solve this problem, a systematic evaluation of all possible combinations of the three most intense resonances was carried out. The two less intense resonances were not taken into account in these calculations, since they correspond to minor conformers of the CLaNP molecule. The multiple resonances for each amide results in a spectral overlap of peaks of some residues. However, the slope $\approx$1 observed for the paramagnetically shifted resonan-

Table 1. Magnetic susceptibility tensor parameters.

	$\Delta\chi_{ax}$ [10^{-32} m^3]	$\Delta\chi_{rh}$ [10^{-32} m^3]	x^t–x'[a]	y^t–y'[a]	z^t–z'[a]
tensor A	12.6	14.8	52.3	168.8	128.4
tensor B	10.2	12.3	47.8	57.9	49.5
tensor C	8.0	8.4	31.6	151.3	161.5

[a] Angles in degrees between the x, y, and z axes of the tensors and the arbitrarily defined reference system (see text). The $\Delta\chi$ values and axes orientations are defined in such a way that the relationship $\Delta\chi_{ax}$, $\Delta\chi_{rh} > 0$ always holds.

ces greatly reduces the assignment problem, because it links all resonances from one amide (Figure 4). Assignment of the paramagnetically shifted resonances can thus be achieved based on the assignments of the diamagnetic resonances and is facilitated by the presence of a fraction of unlabelled protein in the sample of the paramagnetic complex. Therefore, a subset of 31 unequivocally assigned amides in this complex, each with three resonances, was selected for tensor determination.

The number of possible combinations increases very rapidly with the number of amides, n (as 6^{n-1}). Thus, subsets of 10 ^{1}H-amide shifts were used to search all possible combinations and to determine the best set of three tensors to describe the PCS. Then, the sets were combined and further optimised by using a genetic algorithm. The best solutions for the three tensors (Figure 5) are given in Table 1.

The observed versus calculated shifts of the 31 amide protons used for tensor determinations are plotted, for each tensor, as solid symbols (●) in Figure 6. Tyr74 is found to be a consistent outlier in the fits between observed and predicted PCS. The reasons for this discrepancy are unclear, but it may reflect a structural difference between crystallographic and solution conditions with regard to the position of the amide proton of this residue. The tensors determined in this way were used to predict the shifts of another 37 residues in the protein to aid in the assignment of their corresponding resonances. These are shown in Figure 6 with open symbols (○).

It could be shown that the solutions obtained for the tensors are not unique. However, the excellent agreement found between the predicted and observed values for the shifted resonances of the 37 amide protons that were not included in the initial analysis clearly demonstrates that the computed tensors are able to describe the paramagnetic effect caused by the three main isomers of the CLaNP molecule.

Distance restraints without tensors: The occurrence of multiple resonances for each amide, caused by the isomers of the probe, offers the possibility to derive distance restraints

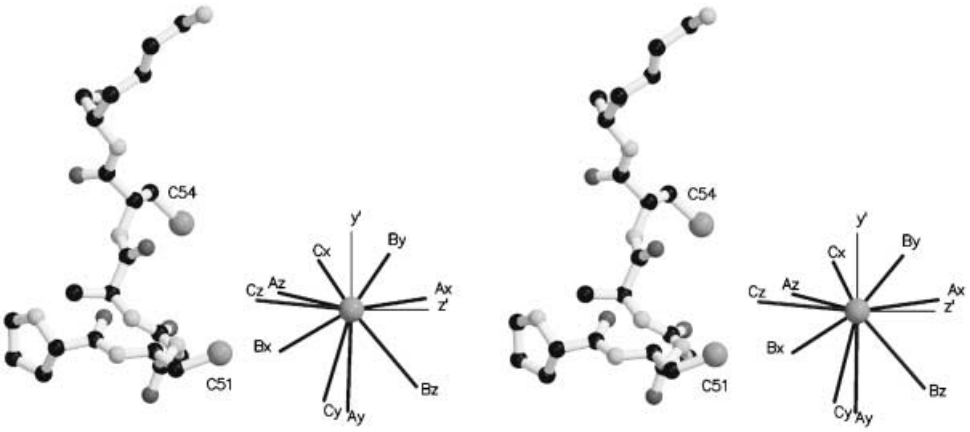

Figure 5. Stereoview of the orientation of the magnetic susceptibility tensors for the three main isomers of the ^{Yb}CLaNP molecule projected onto the Y atom of the ^{Y}CLaNP–dCPsaz complex structure. The axes of tensors A, B and C (see also Table 1) are indicated, as well as the reference frame (y′ and z′; x′ is pointing towards the back). Carbon, oxygen and nitrogen atoms are shown in black, dark grey and light grey, respectively. The sulphur atoms of the Cys residues and the Y atom are shown as large spheres.

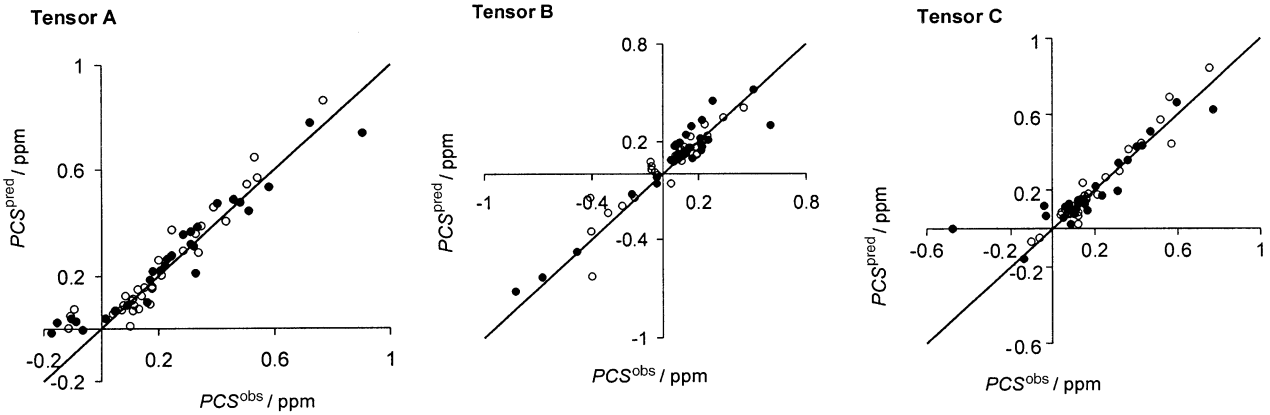

Figure 6. Plots of observed versus predicted PCS for 31 amide protons used for tensor determination (●) and an additional 37 amide protons assigned with the aid of the determined tensors (○). The solid lines represent $PCS^{pred} = PCS^{obs}$.

without using the magnetic anisotropy tensors. This is illustrated in Figure 7. The largest positive PCS of the three intense resonances observed for each amide proton is plotted against the distance between the proton and the Yb atom. The best fit with the inverse cubic power of the distance is shown as a solid line. The dashed line indicates the theoretical maximum (based on the largest tensor in Table 1) and the lower 95% prediction interval of the fit is shown by the dotted line. Given an observed PCS, a target distance and distance range can be read off by using the solid, dotted and dashed lines, respectively. For example, a maximum PCS of 0.6 ppm yields a target of 17.5 Å and the distance range of 16–22 Å. Like with NOE-based distance restraints, the low precision of the target needs not be problematic, provided sufficient restraints can be observed. Such PCS distance restraints could be useful in docking calculations of two proteins, as was demonstrated before on the complex of plastocyanin and cytochrome *f*.[38,39] For this purpose, the position of the lanthanide can be modelled on the basis of the crystal structure of ^{Y}CLaNP–dCPsaz, with the lanthanide in a plane with two Cys sulphur atoms at 7.3 Å from the two S^{γ}(Cys) atoms with Ln/S^{γ}(Cys)/C^{β}(Cys) angles at 129°.

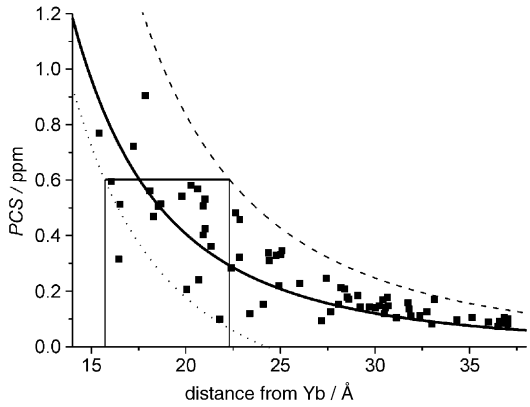

Figure 7. Estimation of distance restraints based on largest observed PCS for 68 amide protons. The lines represent the best fit with the inverse cubic power of the distance (——), the theoretical maximum (-----) and the lower 95% prediction interval of the fit (•••••). A maximum PCS of 0.6 ppm yields a target distance of 17.5 Å and distance range of 16–22 Å.

Conclusion

We have designed a paramagnetic molecule, CLaNP, which can be attached specifically to a protein surface by using two engineered Cys residues at 8–10 Å distance. This bidentate mode of attachment sufficiently reduces the mobility of the Ln^{3+} ion relative to the protein to allow measurement of PCS. Both NMR chemical shift analysis and X-ray diffraction indicate that the structure of the protein is essentially unaffected by the attachment of the probe. This makes CLaNP the first general NMR probe designed to generate PCS. In the current example, shifts induced by Yb are observed up to 40 Å away from the metal and lanthanides that induce larger PCS, such as Dy, are expected to extend this range even further. Conversely, the shift range can be tuned down by choosing lanthanides with a lower ability to shift. The occurrence of several isomers of the CLaNP is still a drawback presented by this molecule and work to synthesise new probes with less or no isomers is in progress. We have demonstrated, however, that the magnetic anisotropy tensors of the different isomers can be calculated, despite the degeneracy problem. We showed that the multiple resonances for each amide also offer an easy and fast way to obtain distance restraints.

The CLaNP molecule can also cause spontaneous alignment at high magnetic fields, which provides a convenient way to obtain residual dipolar couplings. The multiple isomers could be beneficial in this case, because they result in multiple alignment orientations of a protein in one sample and in one experiment.

Finally, we are investigating the potential of CLaNP as a relaxation agent. It can be expected that CLaNP containing gadolinium can provide distance restraints on the basis of relaxation. Contrary to other lanthanides, Gd^{3+} has a long electronic relaxation time causing strong relaxation effects. Due to the isotropic nature of paramagnetic relaxation, it is expected that isomeric forms of the probe will be irrelevant to this application.

The CLaNP molecule offers the possibility of attaching magnetically anisotropic Ln^{3+} ions to any protein, irrespective of its native ability to bind lanthanides. Thus, CLaNP represents a generally useful tool to provide variable-range distance restraints in structure determination of proteins and protein complexes, in the same way as is now standard in paramagnetic metal proteins.

Experimental Section

Synthesis of the probe molecule: To prepare the bisanhydride form of DTPA, a suspension of DTPA (49 g, 125 mmol) in acetic anhydride (53 mL, 560 mmol) and dry pyridine (62 mL, 770 mmol) was heated at 65 °C, while stirring for 24 h. After cooling to room temperature, the precipitate formed was filtered off and then washed with diethyl ether. After drying under vacuum, pure DTPA-bis(anhydride) was obtained (43 g, 120 mmol, 96%). ^{13}C NMR ([D$_6$]DMSO): $\delta = 50.78$, 51.72, 52.71, 54.73, 165.88, 171.82 ppm. The sodium salt of the bis(MTS) derivative of DTPA was prepared by dissolving the HBr salt of MTS (25.9 mg, 0.11 mmol) (Toronto Research Chemicals) in DMSO (0.5 mL). In this solution, DTPA-bisanhydride (19.3 mg, 0.054 mmol) was dissolved by heating. After cooling to room temperature, this solution was washed with DMSO (0.1 mL) into *N*-methylmorpholine (33.2 mg). After standing for 2 h at room temperature and a night at −20 °C, the solvent and the *N*-methylmorpholine were evaporated under vacuum (0.1 mbar). The glassy solid was dissolved in a solution of Na_2CO_3 (23.8 mg) in water (1.5 mL). The solution was filtered and then freeze-dried. The Ln-containing probe molecule was prepared by mixing equimolar amounts of the lanthanide ion and the bis(MTS) derivative of DTPA.

Protein expression and purification: The pseudoazurin gene from *Alcaligenes faecalis* S-6 was amplified from the pUB1 vector kindly provided by Dr. Makoto Nishiyama at the University of Tokyo by using the forward primer, ′A/AG-CTT G/CT-AGC GAA AAT ATC GAA GTT CAT ATG CT 3′ and the reverse primer, 5′CAT C/TC-GAG TCA TTT AGC GCT GGC GAT GAC 3′. The PCR product was digested by using the restriction endonucleases *Hin*dIII and *Xho*I and cloned into the transfer vector pBluescript® II SK- (Stratagene, La Jolla, CA). This was followed by a second cloning step into the *Nhe*I and *Xho*I sites of the bacterial expression vector pET24c (Novagen Inc). From DNA sequence analysis of

the pEPsaz plasmid obtained in this way, it was confirmed that no mutations were generated during PCR amplification. For cloning purposes, extra Ala and Ser residues were introduced at the −1 and 0 positions, respectively. The modifications encoding the E51C and E54C mutations were introduced in a single step by site-directed mutagenesis following a procedure based on Stratagene's ExSite® PCR-Based Site-Directed Mutagenesis Kit and by using the pEPsaz plasmid as a template. High level expression yielding soluble protein was observed in rich medium, while the mutant protein was found in inclusion bodies when cultured on minimal medium. Therefore, uniformly ^{15}N-labelled dCPsaz was obtained from culturing *Escherichia coli* BL21(DE3) cells in E. Coli-OD4 N medium (from Silantes GmbH), containing kanamycin (50 μg mL^{-1}). Cells were grown at 37 °C to $OD_{600} \approx 0.7$, induced with IPTG (0.5 mM) and allowed to grow for a further 5 h at 30 °C. Following harvesting by centrifugation, the cells were resuspended in a buffer (100 mM Tris-HCl, pH 7.2 containing 0.5 M NaCl and 1 mM $ZnCl_2$) and disrupted in a French pressure cell in the presence of phenylmethylsulfonyl fluoride (PMSF) (1 mM) and DNase I (50 μg mL^{-1}). After dialysis against MES (10 mM pH 6.5) with $ZnCl_2$ (1 mM), dCPsaz was purified by separating the crude extract in a cation exchange carboxymethyl cellulose (CM) column equilibrated with MES (10 mM pH 6.5) and eluted with a gradient of NaCl (0–250 mM) in the same buffer; this was followed by gel filtration on a Superdex G75 column equilibrated with a sodium phosphate buffer (50 mM, pH 7.0) containing NaCl (150 mM). The wild-type ^{15}N(/^{13}C)-Zn-pseudoazurin was purified as described above from cells grown in M9 minimal medium containing $^{15}NH_4Cl$ (and uniformly ^{13}C labelled D-glucose) as the sole nitrogen (and carbon) source(s).

Construction of the protein–CLaNP complex: To substitute zinc for copper, wild-type pseudoazurin and E51C/E54C-pseudoazurin samples were prepared by the addition of cyanide (100 mM) to a concentrated protein solution followed by immediate removal of the CN^- ion on a desalting Superdex G25 column equilibrated with MES (10 mM, pH 6.5) containing $ZnCl_2$ (1 mM). Excess Zn was removed on a desalting PD10 column equilibrated with sodium phosphate (20 mM, pH 7.0). Before the reaction between dCPsaz and the lanthanide probe, protein dimers were dissociated by incubating the protein with DTT (3.5 mM) for an hour at room temperature; this was followed by DTT removal on a desalting PD10 column, equilibrated with degassed sodium phosphate (20 mM pH 7.0). Two molar equivalents of the probe molecule were added dropwise to a protein solution (10 mL) in a sodium phosphate buffer (20 mM, pH 7.0). The reaction proceeded overnight at 4 °C under semi-anaerobic conditions. Monomeric protein was separated from the reaction mixture and dimer formed on a Superdex G75 column equilibrated with sodium phosphate buffer (50 mM, pH 7.0), which contained NaCl (150 mM). Approximately 60 % of this fraction contains the complex formed between the protein and the lanthanide probe (dCPsaz–^{Ln}CLaNP). The remaining, free, protein was used as an internal diamagnetic control to facilitate the assignment of [^{15}N,^{1}H]-HSQC peaks.

NMR spectroscopy: NMR samples of wild-type pseudoazurin contained protein (1.5 mM) in sodium phosphate (20 mM), pH 7.0, 6 % (v/v) D_2O. Samples of dCPsaz–^{Ln}CLaNP contained protein (0.5–1 mM) in the same buffer. Measurements were performed on a Bruker Avance DMX600 spectrometer operating at 303 K. [^{15}N,^{1}H]-HSQC spectra were obtained with spectral widths of 32 ppm (^{15}N) and 12.6 ppm (^{1}H). Data processing was performed in AZARA (available from ftp://ftp.bio.cam.ac.uk/pub/azara). Assignments of the ^{1}H and ^{15}N amide nuclei in Zn-pseudoazurin (see Table S1 in the Supporting Information) were obtained with a ^{15}N,^{13}C-labelled sample by using an HNCACB experiment, in conjunction with previous assignments of Cu^+-containing pseudoazurin (results to be published elsewhere). Spectra were analysed by using the program 'Ansig for Windows'.[40]

Mass spectrometry: Protein mass determination was performed with a Q-TOF1 mass spectrometer (MS) (Micromass, Manchester, UK), equipped with an on-line nano-electrospray source. The samples were diluted to a concentration of 3 pmol μL^{-1} in water/methanol/acetic acid 50:50:1 (v/v/v) and introduced into the MS by flow injection analysis. Mass spectra were recorded from *m/z* 50–2000. The protein mass was calculated from the protein envelope by deconvolution with the MaxEnt software.

Crystallisation and structure determination: Crystals of the ^{Y}ClaNP–dCPsaz complex were grown at 18 °C in a mother liquor, which consisted of ammonium sulphate (2.4 M) and sodium phosphate (50 mM) at pH 6.5. These conditions resulted in the growth of twinned crystals that were characterised as belonging to space group $P6_5$ with cell dimensions $a = b = 42.7$ Å, $c = 116.9$ Å. The crystals were soaked in a mother liquor with glycerol (30 %) as a cryoprotectant and were looped directly into a cryostream at 100 K. X-ray data were collected to 2.0 Å resolution on a MAR345 image plate system and a home laboratory X-ray source. The data were processed with DENZO[41] and corrected for twinning by using the program Detwin[42] and a 0.30 twin fraction. The structure of wild-type Psaz from *A. faecalis*[43] was used as the starting model for molecular replacement with the program MOLREP[44] following the removal of the copper and all solvent atoms. Eight percent of the data (573 reflections) were set aside to calculate the free *R* factor. Refinement was accomplished by using CNS (for water picking)[45] and Refmac5.[46] The final structure includes the recombinant Psaz peptide (residues Ala-1 to Ala-120), a Zn atom at the type I site and the complete probe, this resulted in an R_{work} and an R_{free} of 19.3 % and 23.2 %, respectively. Statistics of the data processing and structure refinement are presented in Table 2.

Table 2. Crystallographic data collection and refinement statistics

Data collection	
resolution [Å]	2.0 (2.07–2.0)
R merge	0.057 (0.343)
{*I*}/{σ(*I*)}	25.3 (5.4)
completeness [%]	99.2 (99.1)
unique reflections collected	8144 (2616)
refinement statistics	
working *R* factor	0.193
free *R* factor	0.232
rmsd bond length [Å]	0.016
overall *B* factor [Å]^2	28.5

Determination of the magnetic susceptibility tensors: Experimental pseudocontact shifts (δ^{pc}) were obtained from the difference between the shifts observed for the paramagnetic (^{Yb}CLaNP–dCPsaz) and the diamagnetic (^{Y}CLaNP–dCPsaz) complexes, $\delta^{pc} = \delta_{obs}$ (paramagnetic complex) $-\delta_{obs}$ (diamagnetic complex). The pseudocontact shift experienced by a nucleus i in a paramagnetic molecule, δ_i^{pc}, is described by Equation 1,

$$\delta_i^{pc} = \frac{1}{12\pi r_i^3}\left[\Delta\chi_{ax}(3\cos^2\theta_i - 1) + \frac{3}{2}\Delta\chi_{rh}(\sin^2\theta_i \cos 2\Omega_i)\right] \quad (1)$$

in which r_i, θ_i and Ω_i are polar coordinates of nucleus i with respect to the principal axes system of the magnetic susceptibility tensor. The axial and rhombic anisotropies, $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$, respectively, are defined as $\Delta\chi_{ax} = \chi_{zz} - 0.5(\chi_{xx} + \chi_{yy})$ and $\Delta\chi_{rh} = \chi_{xx} - \chi_{yy}$, in which χ_{xx}, χ_{yy} and χ_{zz} are the principal components of the magnetic susceptibility tensor. The Euler rotation matrix $\Gamma(\alpha,\beta,\gamma)$ converts the arbitrarily defined, metal-centred, molecular coordinate system (the reference system), with polar (Cartesian) coordinates r,θ',Ω' (x′, y′, z′), into the magnetic coordinate system r,θ^t,Ω^t (x^t, y^t, z^t). The arbitrary coordinates, x′, y′, z′, were determined by using the X-ray crystal structure of the ^{Y}CLaNP–dCPsaz complex (PDB entry 1PY0). For this analysis the program MOLMOL[47] was used to add proton coordinates. The reference frame was defined as having its origin on the lanthanide atom and applying the Euler rotation matrix in Table 3 to the coordinates of the atoms in the structure of the ^{Y}CLaNP–dCPsaz complex.

In the ^{Yb}CLaNP–dCPsaz complex, 31 amides were assigned unequivocally with each amide showing three equally intense resonances (A, B and C) and corresponding pseudocontact shifts, $\delta_{i,j}^{pc}$ (j = A,B,C), which represent three isomers of the CLaNP molecule.

Table 3. Euler rotation matrix values for definition of the x′, y′, z′ axes system

0.134	0.051	−0.990
0.935	0.325	0.144
0.329	−0.944	−0.004

By using two subsets of 10 amides each, all permutations of A, B and C were used to calculate magnetic anisotropy tensors by means of a modified version of the programme FANTASIAN.[48] The tensors were sorted according to their capability to predict the input data and the best solutions were used as starting values for optimisation by using all 31 data points with an evolutionary algorithm called differential evolution.[49,50] Details of this approach are given in the Supporting Information. The software used in the calculations is available upon request.

The final tensors were used to predict the shifts (A, B, C) of another 37 amides, which helped to assign these amides unequivocally. These data points are indicated in Figure 6 with open symbols (○).

Acknowledgements

We would like to thank Prof. G. W. Canters for many useful discussions and Prof. C. Luchinat for critical reading of the manuscript. Prof. I. Bertini is acknowledged for proving the source code of FANTASIAN. This work was supported by the Research Training Network 'Transient' (contract number HPRN-CT-1999-00095) and a National Science and Engineering Research Council Discovery Grant to MEPM. MP acknowledges the financial support provided through the European Community's Human Potential Programme under contract HPRN-CT-1999-00095, TRANSIENT, and a Marie Curie Fellowship (contract number HPMF-CT-2000–00928). MEPM is supported by a Canadian Institutes of Health Research Scholarship.

[1] I. Bertini, C. Luchinat, G. Parigi, *Solution NMR of Paramagnetic Molecules*, **2001,** Elsevier, Amsterdam.
[2] I. Bertini, A. Rosato, P. Turano, *Pure Appl. Chem.* **1999**, *71*, 1717–1725.
[3] L. Banci, I. Bertini, L. D. Eltis, I. C. Felli, D. H. W. Kastrau, C. Luchinat, M. Piccioli, R. Pierattelli, M. Smith, *Eur. J. Biochem.* **1994**, *225*, 715–725.
[4] I. Bertini, A. Donaire, B. Jiménez, C. Luchinat, G. Parigi, M. Piccioli, L. Poggi, *J. Biomol. NMR* **2001**, *21*, 85–98.
[5] I. Bertini, M. B. L. Janik, Y. M. Lee, C. Luchinat, A. Rosato, *J. Am. Chem. Soc.* **2001**, *123*, 4181–4188.
[6] M. Allegrozzi, I. Bertini, M. B. L. Janik, Y. M. Lee, G. H. Lin, C. Luchinat, *J. Am. Chem. Soc.* **2000**, *122*, 4154–4161.
[7] J. R. Tolman, J. M. Flanagan, M. A. Kennedy, J. H. Prestegard, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 9279–9283.
[8] L. R. Dick, C. F. G. C. Geraldes, A. D. Sherry, C. W. Gray, D. M. Gray, *Biochemistry* **1989**, *28*, 7896–7904.
[9] G. Veglia, S. J. Opella, *J. Am. Chem. Soc.* **2000**, *122*, 11733–11734.
[10] M. Sattler, S. W. Fesik, *J. Am. Chem. Soc.* **1997**, *119*, 7885–7886.
[11] E. Liepinsh, M. Baryshev, A. Sharipo, M. Ingelman-Sundberg, G. Otting, S. Mkrtchian, *Structure* **2001**, *9*, 457–471.
[12] S. Arumugam, C. L. Hemme, N. Yoshida, K. Suzuki, H. Nagase, M. Bejanskii, B. Wu, S. R. Van Doren, *Biochemistry* **1998**, *37*, 9650–9657.
[13] S. Aime, N. D'Amelio, M. Fragai, Y. M. Lee, C. Luchinat, E. Terreno, G. Valensin, *J. Biol. Inorg. Chem.* **2002**, *7*, 617–622.
[14] G. Pintacuda, G. Otting, *J. Am. Chem. Soc.* **2002**, *124*, 372–373.
[15] A. M. Petros, L. Mueller, K. D. Kopple, *Biochemistry* **1990**, *29*, 10041–10048.
[16] M. Scarselli, A. Bernini, C. Segoni, H. Molinari, G. Esposito, A. M. Lesk, F. Laschi, P. Temussi, N. Niccolai, *J. Biomol. NMR* **1999**, *15*, 125–133.
[17] T. Yuan, H. Ouyang, H. J. Vogel, *J. Biol. Chem.* **1999**, *274*, 8411–8420.
[18] S. U. Dunham, S. U. Dunham, C. J. Turner, S. J. Lippard, *J.Am.-Chem.Soc.* **1998**, *120*, 5395–5406.
[19] A. Ramos, G. Varani, *J. Am. Chem. Soc.* **1998**, *120*, 10992–10993.
[20] J. L. Battiste, G. Wagner, *Biochemistry* **2000**, *39*, 5355–5365.
[21] V. Gaponenko, J. W. Howarth, L. Columbus, G. Gasmi-Seabrook, J. Yuan, W. L. Hubbell, P. R. Rosevear, *Protein Sci.* **2000**, *9*, 302–309.
[22] A. Lewin, J. P. Hill, R. Boetzel, T. Georgiuo, R. James, C. Kleanthous, G. R. Moore, *Inorg. Chim. Acta* **2002**, *331*, 123–130.
[23] K. Tu, M. Gochin, *J. Am. Chem. Soc.* **1999**, *121*, 9276–9285.
[24] V. Gaponenko, A. S. Altieri, J. Li, R. A. Byrd, *J. Biomol. NMR* **2002**, *24*, 143–148.
[25] J. Iwahara, D. E. Anderson, E. C. Murphy, G. M. Clore, *J. Am. Chem. Soc.* **2003**, *125*, 6634–6635.
[26] J. T. Welch, W. R. Kearney, S. J. Franklin, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3725–3730.
[27] I. Bertini, M. B. L. Janik, Y. M. Lee, C. Luchinat, A. Rosato, *J. Am. Chem. Soc.* **2001**, *123*, 4181–4188.
[28] R. Barbieri, I. Bertini, G. Cavallaro, Y. M. Lee, C. Luchinat, A. Rosato, *J. Am. Chem. Soc.* **2002**, *124*, 5581–5587.
[29] M. Allegrozzi, I. Bertini, S. N. Choi, Y. M. Lee, C. Luchinat, *Eur. J. Inorg. Chem.* **2002**, 2121–2127.
[30] V. Gaponenko, A. Dvoretsky, C. Walsby, B. M. Hoffman, P. R. Rosevear, *Biochemistry* **2000**, *39*, 15217–15224.
[31] J. Feeney, B. Birdsall, A. F. Bradbury, R. R. Biekofsky, P. M. Bayley, *J. Biomol. NMR* **2001**, *21*, 41–48.
[32] C. Ma, S. J. Opella, *J. Magn. Reson.* **2000**, *146*, 381–384.
[33] J. Wohnert, K. J. Franz, M. Nitz, B. Imperiali, H. Schwalbe, *J. Am. Chem. Soc.* **2003**, *125*, 13338–13339.
[34] A. Dvoretsky, V. Gaponenko, P. R. Rosevear, *FEBS Lett.* **2002**, *528*, 189–192.
[35] C. F. G. C. Geraldes, A. M. Urbano, M. A. Hoefnagel, J. A. Peters, *Inorg. Chem.* **1993**, *32*, 2426–2432.
[36] H. Nar, R. Huber, A. Messerschmidt, A. C. Filippou, M. Barth, M. Jaquinod, M. Vandekamp, G. W. Canters, *Eur. J. Biochem.* **1992**, *205*, 1123–1129.
[37] M. Gochin, H. Roder, *Protein Sci.* **1995**, *4*, 296–305.
[38] M. Ubbink, M. Ejdeback, B. G. Karlsson, D. S. Bendall, *Structure* **1998**, *6*, 323–335.
[39] P. B. Crowley, G. Otting, B. G. Schlarb-Ridley, G. W. Canters, M. Ubbink, *J. Am. Chem. Soc.* **2001**, *123*, 10444–10453.
[40] M. Helgstrand, P. Kraulis, P. Allard, T. Härd, *J. Biomol. NMR* **2000**, *18*, 329–336.
[41] Z. Otwinowski, W. Minor, *Methods Enzymol.* **1997**, *276*, 307–326.
[42] T. O. Yeates, *Methods Enzymol.* **1997**, *276*, 344–358.
[43] K. Petratos, Z. Dauter, K. S. Wilson, *Acta Crystallogr. Sect.B.* **1988**, *44*, 628–636.
[44] A. Vagin, A. Teplyakov, *J. Appl. Crystallogr.* **1997**, *30*, 1022–1025.
[45] A. T. Brunger, P. D. Adams, G. M. Clore, W. L. Delano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, G. L. Warren, *Acta Crystallogr. Sect. D* **1998**, *54*, 905–921.
[46] G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallogr. Sect. D* **1997**, *53* 240–255.
[47] R. Koradi, M. Billeter, K. Wüthrich, *J. Mol. Graphics* **1996**, *14* 51–55.
[48] L. Banci, I. Bertini, G. G. Savellini, A. Romagnoli, P. Turano, M. A. Cremonini, C. Luchinat, H. B. Gray, *Proteins: Struct. Funct. Genet.* **1997**, *29*, 68–76.
[49] R. Storn, *IEEE Trans. Evol. Comput.* **1999**, *3* 22–34.
[50] R. Storn, K. Price, *J.Global Optimiz.* **1997**, *11* 341–359.
[51] R. J. Read, *Acta Crystallogr. Sect. A* **1986**, *42* 140–149.

Received: December 12, 2003

Published online: May 6, 2004

